

Modeling Group Interactions via Open Data Sources (AFOSR FA9550-08-1-0132)

Project Description: This project addresses an emerging challenge of developing a computational understanding of online groups and their interactions. By investigating the interactions between the groups and the social, economical, and physical forces acting on them, this project endeavors to achieve a better understanding of the operational environment by extracting pertinent values from groups of special interests. Online data sources contain massive amounts of data. The state-of-art search engines are designed to help general query-specific search and not suitable for finding disconnected online groups. The scientific and technical merits of the proposed research are demonstrated in (1) formulating novel research problems of searching disconnected online groups, (2) developing innovative mathematical and statistical models and efficient algorithms that leverage existing search engines and employ contextual information and meta data, and (3) conducting interdisciplinary scientific research that enables a new kind of search power and complementary research capabilities.

PI: Dr. Huan Liu, CIDSE, ASU; **Point of Contact** at AFRL: Dr. John Salerno

Project summaries: The project develops new research capabilities, broadens the research base in online group research, contributes to improving business intelligence, advancing scientific collaboration, and expediting the development of a new generation of search tools. The project also nurtures graduate students, helps renovate and develop CSE courses, and promotes interdisciplinary research activities via annual workshops and conferences.

Accomplishments

- 3 Dissertations, 4 Books, 6 Book chapters, 1 Encyclopedia article, 12 Journal papers, 18 fully refereed Conference papers, 3 Workshop papers, 5 Technical reports, 3 Invited talks, 2 Tutorials, 1 new course, 2 Journal special issues during the funded period.
- 2 Conferences: International Conference on Social Computing, Behavioral Modeling, & Prediction (SBP 2010/2011)
- 2 Workshops: International Workshop on Social Computing, Behavioral Modeling, & Prediction (SBP 2008/2009)

Current and Former Student Members

- Graduated Students
 - Nitin Agarwal, PhD, Assistant professor, University of Arkansas at Little Rock
 - Sai Moturu PhD, Postdoctoral Researcher, MIT Media Lab
 - Lei Tang, PhD, Research Scientist, Yahoo! Labs
- Current Students
 - Xufei Wang, PhD student

20120918079

A list of accomplishments during the funded period (03/2008 – 08/2011)

Dissertations

1. Lei Tang. Ph.D. dissertation on "Learning with Large-Scale Social Media Networks". School of Computing and Informatics, Arizona State University, Summer 2010.
2. Sai Moturu. Ph.D. dissertation on "Quantifying the Trustworthiness of User-Generated Social Media Content". School of Computing and Informatics, Arizona State University, Summer 2009.
3. Nitin Agarwal. Ph.D. dissertation on "Social Computing in Blogosphere". School of Computing and Informatics, Arizona State University, Summer 2009.

Journal papers

1. Nitin Agarwal, Huan Liu, Lei Tang, and Philip Yu. "Modeling Blogger Influence in a Community", Social Network Analysis and Mining, Springer, forthcoming
2. Lei Tang, Xufei Wang and Huan Liu. "Community Detection in Multi-Dimensional Networks", Journal Data Mining and Knowledge Discovery (DMKD), Springer, August, 2011.
3. Lei Tang, Xufei Wang and Huan Liu, "Group Profiling for Understanding Social Structures", ACM Transactions on Intelligent Systems and Technology (ACM TIST), accepted.
4. Geoffrey Barbier, Lei Tang and Huan Liu. "Understanding Online Groups through Social Media". WIREs Data Mining and Knowledge Discovery, accepted.
5. Lei Tang, Xufei Wang and Huan Liu. "Scalable Learning of Collective Behavior", IEEE Transaction of Knowledge and Data Engineering (IEEE TKDE), accepted.
6. Lei Tang, Huan Liu and Jianping Zhang. "Identifying Evolving Groups in Dynamic Multi-Mode Networks". In IEEE Transactions on Knowledge and Data Engineering (TKDE), accepted.
7. Sai Moturu and Huan Liu. "Quantifying the Trustworthiness of Social Media Content", Journal of Distributed and Parallel Databases, Springer, Volume 29, January 4, 2011.
8. Lei Tang and Huan Liu. "Leveraging Social Media Networks for Classification". Journal of Data Mining and Knowledge Discovery (DMKD), Springer, January 14, 2011.
9. Lei Tang and Huan Liu. "Toward Predicting Collective Behavior via Social Dimension Extraction". In IEEE Intelligent Systems, 2010.
10. Nitin Agarwal, Magdiel Galan, Huan Liu, and Shankaraa Subramanya. "WisColl: Collective Wisdom based Blog Clustering". Journal of Information Science: Special Issue on Collective Intelligence (INS-CI), 180(1):39-61. 2010.
11. Nitin Agarwal and Huan Liu. "Blogosphere: Research Issues, Tools, and Applications". SIGKDD Explorations, 10(1): pp. 18-31, July 2008.
12. Lei Tang, Huan Liu, Jianping Zhang, Nitin Agarwal and John J. Salerno. "Topic Taxonomy Adaptation for Group Profiling", ACM Transactions on Knowledge Discovery from Data, TKDD 1(4), 2008.

Conference papers

1. Xia Hu, Lei Tang, and Huan Liu. "Enhancing Accessibility of Microblogging Messages Using Semantic Knowledge", the 20th ACM Conference on Information and Knowledge Management (CIKM2011), October 24-28, 2011. Glasgow, UK
2. Xufei Wang, Huan Liu, and Wei Fan. "Connecting Users with Similar Interests via Tag Network Inference", the 20th ACM Conference on Information and Knowledge Management (CIKM2011), October 24-28, 2011. Glasgow, UK.
3. Pritam Gundecha, Geoffrey Barbier, and Huan Liu. "Exploiting Vulnerability to Secure User Privacy on a Social Networking Site", the 17th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2011), August 21-24, 2011, San Diego, CA.
4. Shamanth Kumar, Reza Zafarani, and Huan Liu. "Understanding User Migration Patterns in Social Media ", the special track on AI and the Web (AIW) at the Twenty-Fifth AAAI Conference on Artificial Intelligence (AAAI-11). August 7 - 11, 2011, in San Francisco, CA. Huiji Gao, Xufei Wang, Geoffrey Barbier, and Huan Liu. Promoting Coordination for Disaster Relief - From Crowdsourcing to Coordination. SBP 2011: 197-204.
5. Geoffrey Barbier, and Huan Liu. Information Provenance in Social Media. SBP 2011: 276-283.
6. Xufei Wang, Lei Tang and Huan Liu. "Discovering Overlapping Groups in Social Media", the 10th IEEE International Conference on Data Mining series (ICDM 2010), December 14 - 17, 2010, Sydney, Australia.
7. Lei Tang, Geoffrey Barbier, Huan Liu and Jianping Zhang, "A Social Network Analysis Approach to Detecting Suspicious Online Financial Activities", International Conference on Social Computing, Behavioral Modeling and Prediction (SBP10), 2010.
8. Nitin Agarwal, Huan Liu, Shankara Subramanya, John Salerno, and Philip Yu. "Connecting Sparsely Distributed Similar Bloggers", IEEE International Conference on Data Mining (ICDM'09), Dec. 6-9, 2009. Miami, USA.
9. Lei Tang, Xufei Wang and Huan Liu, "Uncovering Groups via Heterogeneous Interaction Analysis", IEEE International Conference on Data Mining (ICDM'09), Dec. 6-9, 2009. Miami, USA.
10. Lei Tang and Huan Liu. "Scalable Learning of Collective Behavior based on Sparse Social Dimensions", The 18th ACM Conference on Information and Knowledge Management (CIKM'09), Hong Kong, Nov. 2-6, 2009.
11. Sai T. Moturu, Jian Yang, and Huan Liu. "Quantifying Utility and Trustworthiness for Advice Shared on Online Social Media", Symposium on Social Intelligence and Networking, IEEE International Conference on Social Computing (SocialCom'09), Vancouver, Canada, Aug 29-31, 2009.
12. Sai T. Moturu and Huan Liu, "Evaluating the Trustworthiness of Wikipedia Articles through Quality and Credibility", The 5th International Symposium on Wikis and Open Collaboration, Florida, Oct 25-27, 2009.

13. L. Tang and H. Liu. "Relational Learning via Latent Social Dimension", Poster Paper, The 15th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'09), Paris, France, August 2009.
14. Nitin Agarwal, Huan Liu, Sudheendra Murthy, Arunabha Sen, and Xufei Wang. "A Social Identity Approach to Identify Familiar Strangers in a Social Network", 3rd International AAAI Conference on Weblogs and Social Media (ICWSM09), pp. 2 - 9, San Jose, California, May 17-20, 2009.
15. N. Agarwal, H. Liu, J. Salerno, and Philip S. Yu. "Searching for 'Familiar Strangers' on Blogosphere: Problems and Challenges", NSF Symposium on Next-Generation Data Mining and Cyber-enabled Discovery and Innovation (NGDM'07), Baltimore, MD, October 10-12, 2007.
16. Shankaraa Subramanya, Baoxin Li, and Huan Liu. "Robust Integration of Multiple Information Sources by View Completion", IEEE International Conference on Information Reuse and Integration 2008 (IEEE IRI), pp 398-403, July 13 -15, 2008. Las Vegas, Nevada.
17. Nitin Agarwal, Magdiel Galan, Huan Liu, Shankaraa Subramanya. "Clustering Blogs with Collective Wisdom", 8th International Conference on Web Engineering (ICWE08), Yorktown Heights, NY, July 14-18, 2008.
18. Sai T. Moturu, Huan Liu, and William Johnson. "Trust Evaluation in Health Information on the World Wide Web", IEEE Engineering in Medicine and Biology Conference (EMBC '08), Vancouver, Canada, August 20 - 24, 2008.

Technical Reports and Workshop Papers

1. Huiji Gao, Xufei Wang, Jiliang Tang and Huan Liu. "Network Denoising in Social Media", Technical Report, TR-11-002, School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ 85287, 2011.
2. Lei Tang, Xufei Wang, Huan Liu and Lei Wang, "A Multi-Resolution Approach to Learning with Overlapping Communities", In KDD Workshop on Social Media Analytics, 2010.
3. Lei Tang, Huiji Gao and Huan Liu, "Network Quantification Despite Biased Labels", In KDD Workshop on Mining and Learning with Graphs, 2010.
4. Lei Tang, Xufei Wang and Huan Liu, "Community Detection in Multi-Dimensional Networks", Technical Report, TR10-006, School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, 85287, 2010.
5. Lei Tang, Xufei Wang and Huan Liu, "Understanding Emerging Social Structures --- A Group Profiling Approach", Technical Report, TR-10-002, School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ, 85287, 2010.
6. Lei Tang and Huan Liu, "Uncovering Cross-Dimension Group Structures in Multi-Dimensional Networks", SDM'09 Workshop on Analysis of Dynamic Networks, April, 2009.
7. Nitin Agarwal, Huan Liu, John J. Salerno, Shankaraa Subramanya, and Philip S. Yu. "Familiar Strangers: Connecting Dots on Blogosphere", Technical Report, TR-08-005,

School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ 85287, 2008.

8. Nitin Agarwal, Magdiel Galan, Huan Liu, Shankar Subramanya. "Clustering with Collective Wisdom - A Comparative Study", Technical Report, TR-08-004, School of Computing Informatics, Arizona State University, Tempe, AZ 85287, 2008.

Books and Chapters

1. Geoffrey Barbier and Huan Liu. "Data Mining in Social Media" in Social Network Data Analytics, Editor: Charu C. Aggarwal, Springer, Boston/Dordrecht/London, pp 327 -352. March, 2011.
2. Lei Tang and Huan Liu. Community Detection and Mining in Social Media, Morgan & Claypool Publishers, 2010.
3. Lei Tang and Huan Liu. Understanding Group Structures and Properties in Social Media. In Link Mining: Models, Algorithms and Applications, Editors: Christos Faloutsos, Jiawei Han and Philip S. Yu. In press.
4. Lei Tang and Huan Liu. Graph Mining Applications to Social Network Analysis. In Managing and Mining Graph Data, Editors: Charu Aggarwal and Haixun Wang. Springer, 2010.
5. Huan Liu, John J. Salerno, and Michael J. Young, editors, "Social Computing, Behavioral Modeling, and Prediction", 2009, Springer.
6. Nitin Agarwal and Huan Liu. "Modeling and Data Mining in Blogosphere". Synthesis Lectures on Data Mining and Knowledge Discovery No. 1, Morgan and Claypool Publishers, Robert Grossman (Editor), August 2009. Morgan and Claypool website; Amazon website.
7. Nitin Agarwal and Huan Liu. "Trust in Blogosphere", in Encyclopedia of Database Systems, forthcoming.
8. Huan Liu, John J. Salerno, and Michael J. Young, editors, "Social Computing, Behavioral Modeling, and Prediction", 2008, Springer.
9. Nitin Agarwal, Huan Liu, and Jianping Zhang. "A Study of Friendship Networks and Blogosphere", Handbook of Research on Text and Web Mining, Technologies. Editors: Min Song and Yi-fang Brook Wu. Idea Group Inc, 2008, Chapter XXXVI, pp. 646-669.
10. Nitin Agarwal, Huan Liu, John J. Salerno, and Philip S. Yu. "Searching for 'Familiar Strangers' on Blogosphere", Next Generation of Data Mining, Editors: Hillol Kargupta, Jiawei Han, Philip S. Yu, Rajeev Motwani, and Vipin Kumar. Chapman & Hall/CRC Press, 2008. Chapter 15, pp 295-316.

Invited Talks

1. Invited Tutorial given by Huan Liu at Human Effectiveness Directorate, Air Force Research Labs Workshop on Computational Methods for Complex Data Analysis, Wright-Patterson AFB, Ohio, September, 2011.
2. Invited Lectures given by Huan Liu at Lipari Summer School on Computational Social Science: Computational Social Science: Text and Decisions, Lipari Island, Italy, July 2011.

3. Invited Talk given by Huan Liu at LinkedIn Corporation: Understanding User Behaviors in Social Media, Mountain View, California, USA, 2011.
4. Invited Talk given by Xufei Wang at Beijing University of Aeronautics & Astronautics (BUAA): Connecting Users with Similar Interests via Tag Network Inference, Beijing, China, 2011.

Other Activities

1. A Special Issue on Social Computing in Blogosphere, 2010 IEEE Internet Computing.
2. IEEE SocialCom'09 Tutorial on Social Computing: Community Detection and Behavior Study for Social Computing, 2009.
3. ACM KDD'08 Tutorial on Blogosphere: Research, Tools, and Applications, 2008.
4. Second International Workshop on Social Computing, Behavioral Modeling, and Prediction (SBP09), Phoenix, Arizona. March 31 - April 1, 2009.
5. ACM TKDD Special Issue on Social Computing, Behavioral Modeling, and Prediction, Submission due on September 10, 2008.
6. CSE 494/598: Social Computing and Web Analytics, 2008, 2010
7. CSE 494/598: Social Media Mining, 2011.

Awards

1. ACM Distinguished Scientist (only 41 nominated worldwide) for Computing Innovation in 2010
2. Best Researcher in Computer Science and Engineering, School of Computing, Informatics, and Decision Systems Engineering, Arizona State University

Distinguish Services

1. Steering Committee Co-Chair for International Conference Series on Social Computing, Behavioral-Cultural Modeling, and Prediction (SBP) and a founding co-organizer for SBP
2. Conference Co-Chair for SIAM International Conference on Data Mining, SDM2011 in Phoenix, AZ, and SDM2012 in Anaheim, CA.
3. Conference Program Co-Chair for SIAM International Conference on Data Mining, SDM2009.
4. Advisory Committee Co-Chair for IEEE International Conference on Cyber, Physical and Social Computing (CPSCoM2011), 2011

Contact Information

Huan Liu – huan.liu@asu.edu

REPORT DOCUMENTATION PAGE

THIS DOCUMENT CONTAINS INFORMATION OF INTEREST TO THE PUBLIC. IT IS THE POLICY OF THE NATIONAL ARCHIVES TO MAKE AVAILABLE TO THE PUBLIC INFORMATION THAT IS NOT OTHERWISE AVAILABLE. INFORMATION CONTAINED HEREIN IS UNCLASSIFIED EXCEPT WHERE SHOWN OTHERWISE. DATE OF DECLASSIFICATION: 01-01-2011. AUTHORITY: NATIONAL ARCHIVES. FOR INFORMATION ON THE STATUS OF THIS DOCUMENT, CONTACT THE NATIONAL ARCHIVES. FOR INFORMATION ON THE STATUS OF THIS DOCUMENT, CONTACT THE NATIONAL ARCHIVES. FOR INFORMATION ON THE STATUS OF THIS DOCUMENT, CONTACT THE NATIONAL ARCHIVES.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE 30-08-2011		2. REPORT TYPE Final Report		3. DATES COVERED March 2008 - August 2011	
4. TITLE AND SUBTITLE Modeling Group Interactions via Open Data Sources				5a. CONTRACT NUMBER FA9550-08-1-0132	
				5b. GRANT NUMBER FA9550-08-1-0132	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Huan Liu, Xufei Wang, Lei Tang, Sai Moturu, Nitin Agarwal, John Salerno, Geoff Barbier				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Arizona State University				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Joseph Lyons, Ph.D. and Stephanie A. Bruce, CPIA, CIP Air Force Office of Scientific Research Mathematics, Information and Life Sciences				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-QSR-VA-TR-2012-0100	
12. DISTRIBUTION/AVAILABILITY STATEMENT Publications in forms books, book chapters, journal articles, conference papers, and technical reports. A					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT This proposal addresses an emerging challenge of developing a computational understanding of online groups and their interactions. By investigating the interactions between the groups and the social, economical, and physical forces acting on them, this project endeavors to achieve a better understanding of the operational environment by extracting pertinent values from groups of special interests. Online data sources contain massive amounts of data. The state-of-art search engines are designed to help general query-specific search and not suitable for finding disconnected online groups. The scientific and technical merits of the proposed research are demonstrated in (1) formulating novel research problems of searching disconnected online groups, (2) developing innovative mathematical and statistical models and efficient algorithms that leverage existing search engines and employ contextual information and meta data, and (3) conducting interdisciplinary scientific research that enables a new kind of search power and complementary research capabilities. The potential contributions of the proposed research to the mission of the IIS Air Force are that the theories, algorithms, and					
15. SUBJECT TERMS Social Computing, Behavior Modeling, Prediction, Data Collection, Data Mining in Social Media					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Huan Liu
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER 480 727 7349